

The Optimal Error Bound for the Method of Simultaneous Projections

Simeon Reich^{*†} Rafał Zalas^{*‡}

April 4, 2017

Abstract

In this paper we find the optimal error bound (smallest possible estimate, independent of the starting point) for the linear convergence rate of the simultaneous projection method applied to closed linear subspaces in a real Hilbert space. We achieve this by computing the norm of an error operator which we also express in terms of the Friedrichs number. We compare our estimate with the optimal one provided for the alternating projection method by Kayalar and Weinert (1988). Moreover, we relate our result to the alternating projection formalization of Pierra (1984) in a product space. Finally, we adjust our results to closed affine subspaces and put them in context with recent dichotomy theorems.

Key words and phrases: Linear rate of convergence, optimal error bound, simultaneous projection method.

2010 Mathematics Subject Classification: 41A25, 41A28, 41A44, 41A65.

1 Introduction

Let M_1, \dots, M_r be closed and linear subspaces of a real Hilbert space \mathcal{H} and let $M := \bigcap_{i=1}^r M_i$. By P_C we denote the metric projection onto a nonempty, closed and convex set $C \subseteq \mathcal{H}$. In this paper we consider simultaneous and cyclic projection methods. The following two theorems are well known:

Theorem 1 (von Neumann [21] and Halperin [14]). *For each $x \in \mathcal{H}$,*

$$\lim_{k \rightarrow \infty} \left\| (P_{M_r} \dots P_{M_1})^k(x) - P_M(x) \right\| = 0. \quad (1)$$

Theorem 2 (Lapidus [16] and Reich [20]). *For each $x \in \mathcal{H}$,*

$$\lim_{k \rightarrow \infty} \left\| \left(\frac{1}{r} \sum_{i=1}^r P_{M_i} \right)^k(x) - P_M(x) \right\| = 0. \quad (2)$$

Let $T, T^\infty: \mathcal{H} \rightarrow \mathcal{H}$ be such that $T^k(x)$ converges to $T^\infty(x)$ for every $x \in \mathcal{H}$. Following [6, 10], we say that T^k *converges arbitrarily slowly* to T^∞ if for every sequence $\{a_k\}_{k=0}^\infty \subseteq (0, \infty)$ satisfying $a_k \rightarrow 0$ as $k \rightarrow \infty$, there is $x \in \mathcal{H}$ such that $\|T^k(x) - T^\infty(x)\| \geq a_k$ for every $k = 0, 1, 2, \dots$. We also recall that T^k *converges linearly* to T^∞ if for some $c, f(x) > 0$ and $q \in (0, 1)$, we have

^{*}Department of Mathematics, The Technion - Israel Institute of Technology, 3200003 Haifa, Israel

[†]✉ sreich@tx.technion.ac.il

[‡]✉ zalasrafal@gmail.com

$\|T^k(x) - T^\infty(x)\| \leq cq^k f(x)$ for all $k = 0, 1, 2, \dots$. In this paper T and T^∞ are related to projections onto linear or affine subspaces in which case we use $f(x) = \|x\|$ or $f(x) = \|x - T^\infty(0)\|$, respectively. Note that even in the case of closed linear subspaces the convergence in Theorems 1 and 2 does not have to be linear and moreover, it may indeed be arbitrarily slow. To see this, we now quote a relevant dichotomy result.

Theorem 3 (Bauschke, Deutsch and Hundal [6, 10]). *Let $T := P_{M_r} \dots P_{M_1}$ or $T := \frac{1}{r} \sum_{i=1}^r P_{M_i}$. Then exactly one of the following two statements holds:*

- (i) $\sum_{i=1}^r M_i^\perp$ is closed. Then T^k converges linearly to P_M .
- (ii) $\sum_{i=1}^r M_i^\perp$ is not closed. Then T^k converges arbitrarily slowly to P_M .

Alternative (ii) of the above theorem has recently been extended in the case of the cyclic projection method. Following [3], we say that T^k converges super-polynomially fast to T^∞ on a nonempty set $X \subseteq \mathcal{H}$ if $\|T^k(x) - T^\infty(x)\|/k^{-\alpha} \rightarrow 0$ as $k \rightarrow \infty$ for all $\alpha > 0$ and $x \in X$.

Theorem 4 (Badea and Seifert [3]). *Let $T := P_{M_r} \dots P_{M_1}$. If $\sum_{i=1}^r M_i^\perp$ is not closed, then T^k converges super-polynomially fast to P_M on some dense linear subspace $X \subseteq \mathcal{H}$.*

For more dichotomy and trichotomy results concerning arbitrarily slow convergence we refer the interested reader to [2, 11, 12]. Note that using the above theorems, one can easily see that arbitrarily slow as well as super-polynomially fast convergence may happen only in the infinite dimensional case.

A very natural question related to Theorem 3 (i) is the following one: What is the optimal error bound (smallest possible estimate, independent of x) such that $\|T^k(x) - P_M(x)\| \leq cq^k \|x\|$? This question can be answered by finding the norm of the error operator $T^k - P_M$. In the case of the alternating projection method ($r = 2$), we have, by Aronszajn [1] (inequality), and Kayalar and Weinert [15] (equality),

$$\|(P_{M_2}P_{M_1})^k - P_M\| = \cos(M_1, M_2)^{2k-1}, \quad (3)$$

where by

$$\cos(M_1, M_2) := \sup\{\langle x_1, x_2 \rangle \mid x_i \in M_i \cap (M_1 \cap M_2)^\perp, \|x_i\| \leq 1\} \in [0, 1] \quad (4)$$

we denote the cosine of the Friedrichs angle between the subspaces M_1 and M_2 . In addition, one can show that $\cos(M_1, M_2) < 1$ if and only if $M_1^\perp + M_2^\perp$ is closed; see, for example, [6, Lemma 1.3]. As far as we are aware, for $r > 2$ the exact computation of the error operator norm for both algorithmic operators is still unknown; see, for example, [2, 3, 11, 12, 18, 19]. Even for $r = 2$, the norm $\|(P_{M_1} + P_{M_2})/2 - P_M\|$ seems to be unknown. Note that one could try to find the optimal estimate for the simultaneous projection method by using (3) and the corresponding alternating projection formalization of Pierra [17] in the product space \mathcal{H}^r . This approach, although very natural, is somehow misleading and, when applied directly, provides a weaker result than the optimal one; compare Example 5, Theorem 8 and Example 9 below.

The main contribution of this paper is to extend Theorem 3 (i) in the case of the simultaneous projection method by finding the optimal error bound, that is, by computing the exact value of $\|(\frac{1}{r} \sum_{i=1}^r P_{M_i})^k - P_M\|$ for any $r \geq 2$; see Theorems 8 and 12 below. For $r = 2$ we show that this norm is greater than the norm in (3), which somewhat explains why, in general, the alternating projection method is indeed faster than its simultaneous variant whenever we have linear convergence; see Remark 10.

In addition, we formally extend Theorem 4 for the simultaneous projection method with $T = \frac{1}{r} \sum_{i=1}^r P_{M_i}$ by using the alternating projection formalization in a product space.

Finally, by using a translation argument, we obtain analogous results in the case of affine subspaces; see Corollary 13.

2 Main result

We begin this section with a simple example showing that a direct application of Pierra's alternating projection formalization in a product space indeed leads to linear convergence, but the obtained estimate, as we show in Theorem 8 below, is not optimal.

Example 5 (Alternating projection formalization of Pierra). Let $M_1, \dots, M_r \subseteq \mathcal{H}$ be closed and linear subspaces and $M := \bigcap_{i=1}^r M_i$. Moreover, following Pierra [17], we consider the subsets

$$\mathbf{C} := M_1 \times \dots \times M_r \quad \text{and} \quad \mathbf{D} := \{\mathbf{x} = (x, \dots, x) \mid x \in \mathcal{H}\} \quad (5)$$

of the product space \mathcal{H}^r equipped with the scalar product $\langle \mathbf{x}, \mathbf{y} \rangle := \frac{1}{r} \sum_{i=1}^r \langle x_i, y_i \rangle$, where $\mathbf{x} = (x_1, \dots, x_r)$, $\mathbf{y} = (y_1, \dots, y_r)$, and $x_i, y_i \in \mathcal{H}$, $i = 1, \dots, r$. We recall [9, Fact 3.2, Lemma 3.3] that for any $\mathbf{x} = (x, \dots, x) \in \mathbf{D}$,

$$\|\mathbf{x}\| = \|x\|, \quad (P_{\mathbf{D}} P_{\mathbf{C}})^k(\mathbf{x}) = (T^k(x), \dots, T^k(x)) \quad \text{and} \quad P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x}) = (P_M(x), \dots, P_M(x)), \quad (6)$$

where $T := \frac{1}{r} \sum_{i=1}^r P_{M_i}$. Moreover, by (3), $\|(P_{\mathbf{D}} P_{\mathbf{C}})^k - P_{\mathbf{C} \cap \mathbf{D}}\| = \cos(\mathbf{C}, \mathbf{D})^{2k-1}$. This leads to the following estimate:

$$\begin{aligned} \|T^k(x) - P_M(x)\| &= \|(T^k(x) - P_M(x), \dots, T^k(x) - P_M(x))\| \\ &= \|(P_{\mathbf{D}} P_{\mathbf{C}})^k(\mathbf{x}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x})\| \\ &\leq \cos(\mathbf{C}, \mathbf{D})^{2k-1} \|\mathbf{x}\| \\ &= \cos(\mathbf{C}, \mathbf{D})^{2k-1} \|x\|. \end{aligned} \quad (7)$$

Consequently, if $(P_{\mathbf{D}} P_{\mathbf{C}})^k$ converges linearly to $P_{\mathbf{C} \cap \mathbf{D}}$, that is, if $\cos(\mathbf{C}, \mathbf{D}) < 1$, then T^k converges linearly to P_M . On the other hand, $\mathbf{y} = P_{\mathbf{D}} P_{\mathbf{C}} \mathbf{x} \in \mathbf{D}$ for every $\mathbf{x} \in \mathcal{H}^r$ and we have

$$\|(P_{\mathbf{D}} P_{\mathbf{C}})^k(\mathbf{x}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x})\| = \|(P_{\mathbf{D}} P_{\mathbf{C}})^{k-1}(\mathbf{y}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{y})\| = \|T^{k-1}(y) - P_M(y)\|. \quad (8)$$

Thus $(P_{\mathbf{D}} P_{\mathbf{C}})^k$ converges linearly to $P_{\mathbf{C} \cap \mathbf{D}}$ whenever T^k converges linearly to P_M .

We now prove the following general lemma. A closely related result can be found in [5, Theorem 2.18].

Lemma 6. *Let \mathcal{H} be a real Hilbert space, $T \in B(\mathcal{H})$ and let $M \subseteq F = \text{Fix } T$ be a nonempty, closed and linear subspace. Assume that $P_M = P_M T$ which holds, for example, if $P_F = P_F T$, or T is self-adjoint, or $\|T\| \leq 1$. Then*

$$T^k - P_M = (T - P_M)^k \quad (9)$$

and therefore

$$\|T^k - P_M\| \leq \|T - P_M\|^k. \quad (10)$$

If, in addition, T is normal, that is, $T^*T = TT^*$, then $T - P_M$ is normal too and consequently,

$$\|T^k - P_M\| = \|T - P_M\|^k. \quad (11)$$

Proof. Note that by assumption, $P_M = TP_M = P_M T$ and we can apply the binomial theorem to obtain

$$(T - P_M)^k = \sum_{l=0}^k \binom{k}{l} (-1)^l T^{k-l} P_M^l = T^k + \sum_{l=1}^k \binom{k}{l} (-1)^l P_M = T^k - P_M. \quad (12)$$

Thus (10) follows. The operator P_M is self-adjoint and hence $T - P_M$ is normal. We recall that for any normal $N \in B(\mathcal{H})$, $\|N^k\| = \|N\|^k$; see, for example, [8, Lemma 8.32]. Thus equality (11) follows from (9).

We now show that $P_M = P_M T$ follows from $P_F = P_F T$. Observe that F is a closed linear subspace. Indeed, due to the continuity of F , for every $F \ni x^k \rightarrow x$, we get $x = \lim x^k = \lim T(x^k) = T(x)$. Consequently, since $M \subseteq F$ are both closed linear subspaces, we have $P_M = P_M P_F$; see [8, Lemma 9.2]. This implies that $P_M T = P_M P_F T = P_M P_F = P_M$.

In the next step we show that $P_F = P_F T$ holds for any self-adjoint T . To this end, we recall that by the characterization of the orthogonal projection [8, Theorem 4.9], $y = P_F(x)$ if and only if $y \in F$ and $\langle x - y, z \rangle = 0$ for every $z \in F$. Now note that $P_F T(x) \in F$ and moreover,

$$\langle x - P_F T(x), z \rangle = \langle x - P_F T(x), T(z) \rangle = \langle T(I - P_F)T(x), z \rangle = \langle (I - P_F)T(x), z \rangle = 0, \quad (13)$$

which completes this part of the proof.

Finally, we show that when $\|T\| \leq 1$, then the identity $P_F = P_F T$ also holds. In this case $\|(I + T + \dots + T^{k-1})/k\| \leq 1$ and $\|T^k(x)/k\| \leq \frac{1}{k} \rightarrow 0$. Consequently, by the mean ergodic theorem [13, Corollary VIII.5.4], we have

$$P_F(x) = \lim_k \frac{x + T(x) + \dots + T^{k-1}(x)}{k} = \lim_k \frac{T(x) + T(T(x)) + \dots + T^{k-1}(T(x))}{k} = P_F T(x), \quad (14)$$

which completes the proof. ■

Before formulating our next result, following [2, Definition 3.2], we recall the following generalization of the cosine of the Friedrichs angle for more than two subspaces.

Definition 7. Let $M_1, \dots, M_r \subseteq \mathcal{H}$ be closed linear subspaces and let $M := \bigcap_i M_i$. The *Friedrichs number* is defined by

$$\cos(M_1, \dots, M_r) := \sup \left\{ \frac{1}{r-1} \frac{\sum_{i \neq j} \langle x_i, x_j \rangle}{\sum_{i=1}^r \|x_i\|^2} \mid x_i \in M_i \cap M^\perp \text{ and } \sum_{i=1}^r \|x_i\|^2 \neq 0 \right\}. \quad (15)$$

The above definition coincides in the case of $r = 2$ with (4); see [2, Lemma 3.1]. Moreover, $\cos(M_1, \dots, M_r) \in [0, 1]$. Thus we can indeed extend the notion of the Friedrichs angle $\theta \in [0, \pi/2]$ with the implicit definition $\cos(\theta) = \cos(M_1, \dots, M_r)$.

Theorem 8 (Exact norm value). *Let $M_1, \dots, M_r \subseteq \mathcal{H}$ be closed linear subspaces and let $M := \bigcap_i M_i$. Moreover, let $\mathbf{C}, \mathbf{D} \subseteq \mathcal{H}^r$ be defined as in (5). Then, for every $k = 1, 2, \dots$,*

$$\begin{aligned} \left\| \left(\frac{1}{r} \sum_{i=1}^r P_{M_i} \right)^k - P_M \right\| &= \left\| \frac{1}{r} \sum_{i=1}^r P_{M_i} - P_M \right\|^k \\ &= \left(\frac{r-1}{r} \cos(M_1, \dots, M_r) + \frac{1}{r} \right)^k \\ &= \cos(\mathbf{C}, \mathbf{D})^{2k} \\ &= \|P_{\mathbf{D}} P_{\mathbf{C}} P_{\mathbf{D}} - P_{\mathbf{C} \cap \mathbf{D}}\|^k \\ &= \|(P_{\mathbf{D}} P_{\mathbf{C}} P_{\mathbf{D}})^k - P_{\mathbf{C} \cap \mathbf{D}}\|. \end{aligned} \quad (16)$$

Proof. Note that $T := \frac{1}{r} \sum_{i=1}^r P_{M_i}$ is self-adjoint, $\|T\| \leq 1$ and $\text{Fix } T = M$. Thus the first equality follows from Lemma 6. Moreover, by [2, Proposition 3.6 and Proposition 3.7], we have

$$\left\| \frac{1}{r} \sum_{i=1}^r P_{M_i} - P_M \right\| = \frac{r-1}{r} \cos(M_1, \dots, M_r) + \frac{1}{r} = \cos(\mathbf{C}, \mathbf{D})^2, \quad (17)$$

which yields the second and third equalities. Note that in [2], $\cos(\mathbf{C}, \mathbf{D})$ is defined in \mathcal{H}^r with $\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^r \langle x_i, y_i \rangle$. Since the definition of $\cos(\cdot)$ does not depend on a positive rescaling of the scalar product, equality (17) indeed holds true for \mathcal{H}^r with $\langle \mathbf{x}, \mathbf{y} \rangle := \frac{1}{r} \sum_{i=1}^r \langle x_i, y_i \rangle$, as it is in the framework of Example 5. The last equality again follows from Lemma 6 applied to $\mathbf{T} := P_{\mathbf{D}} P_{\mathbf{C}} P_{\mathbf{D}}$. Furthermore, we see that

$$\begin{aligned} \|P_{\mathbf{D}} P_{\mathbf{C}} P_{\mathbf{D}} - P_{\mathbf{C} \cap \mathbf{D}}\| &= \|P_{\mathbf{D}} P_{\mathbf{C}} P_{\mathbf{C}} P_{\mathbf{D}} - P_{\mathbf{C} \cap \mathbf{D}}\| \\ &= \|(P_{\mathbf{D}} P_{\mathbf{C}} - P_{\mathbf{C} \cap \mathbf{D}})(P_{\mathbf{C}} P_{\mathbf{D}} - P_{\mathbf{C} \cap \mathbf{D}})\| \\ &= \|(P_{\mathbf{D}} P_{\mathbf{C}} - P_{\mathbf{C} \cap \mathbf{D}})(P_{\mathbf{D}} P_{\mathbf{C}} - P_{\mathbf{C} \cap \mathbf{D}})^*\| \\ &= \|P_{\mathbf{D}} P_{\mathbf{C}} - P_{\mathbf{C} \cap \mathbf{D}}\|^2 \\ &= \cos(\mathbf{C}, \mathbf{D})^2, \end{aligned} \quad (18)$$

where the second equality follows from $P_{\mathbf{C} \cap \mathbf{D}} = P_{\mathbf{C} \cap \mathbf{D}} P_{\mathbf{C}} = P_{\mathbf{C} \cap \mathbf{D}} P_{\mathbf{D}}$ and the latter one follows from (3). This completes the proof. ■

Example 9 (Example 5 revisited). In the setting of Example 5, a direct application of Pierra's formalization in a product space leads to an estimate which, in view of Theorem 8, is not the optimal one. The remedy to this problem is to consider $P_{\mathbf{D}}P_{\mathbf{C}}P_{\mathbf{D}}$ instead of $P_{\mathbf{D}}P_{\mathbf{C}}$. Indeed, for any $\mathbf{x} = (x, \dots, x) \in \mathbf{D}$, we have $(P_{\mathbf{D}}P_{\mathbf{C}}P_{\mathbf{D}})^k(\mathbf{x}) = (P_{\mathbf{D}}P_{\mathbf{C}})^k(\mathbf{x})$ and consequently,

$$\|T^k(x) - P_M(x)\| = \|(P_{\mathbf{D}}P_{\mathbf{C}}P_{\mathbf{D}})^k(\mathbf{x}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x})\| \leq \cos(\mathbf{C}, \mathbf{D})^{2k} \|\mathbf{x}\| = \cos(\mathbf{C}, \mathbf{D})^{2k} \|x\|. \quad (19)$$

Although the above inequality recovers the optimal error bound from Theorem 8, it does not explain why this estimate is optimal.

Remark 10 (Two subspaces). Let $M_1, M_2 \subseteq \mathcal{H}$ be closed linear subspaces and let $M := M_1 \cap M_2$. By (3) and Theorem 8,

$$\begin{aligned} \|(P_{M_2}P_{M_1})^k - P_M\| &= \cos(M_1, M_2)^{2k-1} \leq \left(\frac{1}{2} \cos(M_1, M_2) + \frac{1}{2}\right)^{2k-1} \\ &\leq \left(\frac{1}{2} \cos(M_1, M_2) + \frac{1}{2}\right)^k = \left\| \left(\frac{P_{M_1} + P_{M_2}}{2}\right)^k - P_M \right\|, \end{aligned} \quad (20)$$

where the inequalities are strict whenever $\cos(M_1, M_2) < 1$. This somehow explains why, in general, the alternating projection method is indeed faster than its simultaneous variant whenever we have linear convergence. The numerical verification of this observation can be found, for example, in [7, Fig. 1, p. 1071].

Next, we recall the following fact.

Fact 11. Let $M_1, \dots, M_r \subseteq \mathcal{H}$ be closed linear subspaces and let $M := \bigcap_i M_i$. Moreover, let \mathbf{C} and \mathbf{D} be as in Example 5. The following conditions are equivalent:

- (i) $\sum_{i=1}^r M_i^\perp$ is closed.
- (ii) $\cos(M_1, \dots, M_r) < 1$ (subspaces are not aligned).
- (iii) $\left\| \frac{1}{r} \sum_{i=1}^r P_{M_i} - P_M \right\| < 1$.
- (iv) $\|P_{\mathbf{D}}P_{\mathbf{C}} - P_{\mathbf{C} \cap \mathbf{D}}\| < 1$.
- (v) $\cos(\mathbf{C}, \mathbf{D}) < 1$.
- (vi) $\mathbf{C}^\perp + \mathbf{D}^\perp$ is closed in \mathcal{H}^r .

Proof. Note that by [2, Theorem 4.1, (3) \Leftrightarrow (11)], (i) \Leftrightarrow (ii) and (v) \Leftrightarrow (vi). The rest of the proof follows from (17). ■

We can now extend Theorems 3 and 4 in the case of the simultaneous projection method.

Theorem 12 (Dichotomy with optimal error bound). Let $M_1, \dots, M_r \subseteq \mathcal{H}$ be closed linear subspaces, $M := \bigcap_i M_i$ and let $T := \frac{1}{r} \sum_{i=1}^r P_{M_i}$. Then exactly one of the following two statements holds:

(i) $\sum_{i=1}^r M_i^\perp$ is closed. Then T^k converges linearly to P_M and

$$q = \frac{r-1}{r} \cos(M_1, \dots, M_r) + \frac{1}{r} \quad (21)$$

is the smallest possible number, independent of x , in the set of all $q \in (0, 1)$ such that

$$\|T^k(x) - P_M(x)\| \leq q^k \|x\| \quad (22)$$

for all $x \in \mathcal{H}$.

(ii) $\sum_{i=1}^r M_i^\perp$ is not closed. Then T^k converges arbitrarily slowly to P_M . Moreover, there is a dense linear subspace $X \subseteq \mathcal{H}$ on which T^k converges super-polynomially fast to P_M .

Proof. If $\sum_{i=1}^r M_i^\perp$ is closed, then linear convergence and the optimal error bound follow from Theorem 8 and Fact 11.

Assume now that $\sum_{i=1}^r M_i^\perp$ is not closed. The arbitrarily slow convergence is an immediate consequence of Theorem 3 (ii). In order to complete the proof, we have to show that the convergence is super-polynomially fast on some dense linear subspace $X \subseteq \mathcal{H}$. To this end, we apply the alternating projection formalization discussed in Example 5.

Indeed, by Fact 11, $\mathbf{C}^\perp + \mathbf{D}^\perp$ is not closed in \mathcal{H}^r . Consequently, by Theorem 4, there is a dense linear subspace $\mathbf{X} \subseteq \mathcal{H}^r$ on which $(P_{\mathbf{C}}P_{\mathbf{D}})^k$ converges super-polynomially fast to $P_{\mathbf{C} \cap \mathbf{D}}$. Note that since $P_{\mathbf{D}}$ is quasi-nonexpansive, for each $\mathbf{x} \in \mathcal{H}^r$, we have

$$\begin{aligned} \|(P_{\mathbf{D}}P_{\mathbf{C}})^k P_{\mathbf{D}}(\mathbf{x}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x})\| &= \|P_{\mathbf{D}}(P_{\mathbf{C}}P_{\mathbf{D}})^k(\mathbf{x}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x})\| \\ &\leq \|(P_{\mathbf{C}}P_{\mathbf{D}})^k(\mathbf{x}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x})\|. \end{aligned} \quad (23)$$

Consequently, for every $\mathbf{x} \in \mathbf{X}$ and $\alpha > 0$, we have

$$\frac{\|(P_{\mathbf{D}}P_{\mathbf{C}})^k P_{\mathbf{D}}(\mathbf{x}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x})\|}{k^{-\alpha}} \longrightarrow 0 \quad \text{as } k \rightarrow \infty. \quad (24)$$

This implies that $(P_{\mathbf{D}}P_{\mathbf{C}})^k$ converges super-polynomially fast to $P_{\mathbf{C} \cap \mathbf{D}}$ on $P_{\mathbf{D}}(\mathbf{X})$. On the other hand, since $P_{\mathbf{D}}(\mathbf{X}) \subseteq \mathbf{D}$, we can define

$$X := \{x \in \mathcal{H} \mid \mathbf{x} = (x, \dots, x) \in P_{\mathbf{D}}(\mathbf{X})\}. \quad (25)$$

Observe that X is a linear subspace of \mathcal{H} because $P_{\mathbf{D}}(\mathbf{X})$ is linear subspace in \mathcal{H}^r , where the latter fact follows from the linearity of $P_{\mathbf{D}}$. Moreover, by (6), for each $x \in X$ and $\alpha > 0$, we have

$$\frac{\|T^k(x) - P_M(x)\|}{k^{-\alpha}} = \frac{\|(P_{\mathbf{D}}P_{\mathbf{C}})^k(\mathbf{x}) - P_{\mathbf{C} \cap \mathbf{D}}(\mathbf{x})\|}{k^{-\alpha}} \longrightarrow 0 \quad \text{as } k \rightarrow \infty, \quad (26)$$

where $\mathbf{x} = (x, \dots, x)$. Consequently, T^k converges super-polynomially fast to P_M on X . It remains to prove that X is dense in \mathcal{H} or, equivalently, that $P_{\mathbf{D}}(\mathbf{X})$ is dense in \mathbf{D} . Note that the second statement follows from the continuity of the metric projection $P_{\mathbf{D}}$ as we now show. Indeed, let $\mathbf{x} \in \mathbf{D}$. Since \mathbf{X} is dense in \mathcal{H}^r , there is a sequence $\{\mathbf{x}_k\}_{k=0}^\infty \subseteq \mathbf{X}$ such that $\mathbf{x}_k \rightarrow \mathbf{x}$ and the above-mentioned continuity yields $P_{\mathbf{D}}(\mathbf{x}_k) \rightarrow P_{\mathbf{D}}(\mathbf{x}) = \mathbf{x}$. This completes the proof. ■

Corollary 13 (Affine subspaces). *Let $V_1, \dots, V_r \subseteq \mathcal{H}$ be closed affine subspaces and assume that $V := \bigcap_i V_i \neq \emptyset$. Moreover, let $T := \frac{1}{r} \sum_{i=1}^r P_{V_i}$. Then exactly one of the following two statements holds:*

(i) $\sum_{i=1}^r (V_i - V_i)^\perp$ is closed. Then T^k converges linearly to P_V and

$$q = \frac{r-1}{r} \cos(V_1, \dots, V_r) + \frac{1}{r} \quad (27)$$

is the smallest possible number, independent of x , in the set of all $q \in (0, 1)$ such that

$$\|T^k(x) - P_V(x)\| \leq q^k \|x - P_V(0)\| \quad (28)$$

for all $x \in \mathcal{H}$, where $\cos(V_1, \dots, V_r) := \cos(V_1 - V_1, \dots, V_r - V_r)$.

(ii) $\sum_{i=1}^r (V_i - V_i)^\perp$ is not closed. Then T^k converges arbitrarily slowly to P_V . Moreover, there is a dense affine subspace $Y \subseteq \mathcal{H}$ on which T^k converges super-polynomially fast to P_V .

Proof. The proof is based on the translation formula

$$P_C(x) = P_{C-v}(x - v) + v, \quad (29)$$

which holds true for every closed and convex set $C \subseteq \mathcal{H}$, $x, v \in \mathcal{H}$. Note that for any $x \in \mathcal{H}$ and $v \in V$, we have $V_i = M_i + v$ and $V = M + v$, where $M_i := V_i - V_i$ and $M := V - V$ are closed linear subspaces. This holds, in particular, for $v = P_V(0)$ which, when combined with an induction argument, leads to

$$\left(\frac{1}{r} \sum_{i=1}^r P_{V_i} \right)^k (x) - P_V(x) = \left(\frac{1}{r} \sum_{i=1}^r P_{M_i} \right)^k (x - v) - P_M(x - v). \quad (30)$$

If $\sum_i M_i^\perp$ is closed, then, by Theorem 12 (i) and (30), estimate (28) holds with q defined as in (27). If there were another $0 < q < \frac{r-1}{r} \cos(V_1, \dots, V_r) + \frac{1}{r}$ for which (28) holds, then, by (30), this would imply that

$$\left\| \left(\frac{1}{r} \sum_{i=1}^r P_{M_i} \right)^k - P_M \right\| \leq q^k < \left(\frac{r-1}{r} \cos(V_1, \dots, V_r) + \frac{1}{r} \right)^k, \quad (31)$$

which is impossible in view of Theorem 8.

If $\sum_i M_i^\perp$ is not closed then, by Theorem 12 (ii), $(\frac{1}{r} \sum_{i=1}^r P_{M_i})^k$ converges arbitrarily slowly to P_M . This implies, by (30), that $(\frac{1}{r} \sum_{i=1}^r P_{V_i})^k$ also converges arbitrarily slowly to P_V . Moreover, by Theorem 12 (ii), there is a dense subspace $X \subseteq \mathcal{H}$ on which $(\frac{1}{r} \sum_{i=1}^r P_{M_i})^k$ converges super-polynomially fast to P_M . It is easy to see that, by (30), $(\frac{1}{r} \sum_{i=1}^r P_{V_i})^k$ converges super-polynomially fast to P_V on $Y := X + v$, which in its turn is a dense affine subspace of \mathcal{H} . ■

Remark 14. (Cyclic projection method) Let $M_1, \dots, M_r \subseteq \mathcal{H}$ be closed linear subspaces and let $M := \bigcap_i M_i$. By [15, Theorem 1], $P_{M_r \cap M^\perp} \dots P_{M_1 \cap M^\perp} = P_{M_r} \dots P_{M_1} - P_M$. Moreover, by Lemma 6 applied to $T := P_{M_r} \dots P_{M_1}$, we see that

$$(P_{M_r} \dots P_{M_1})^k - P_M = (P_{M_r} \dots P_{M_1} - P_M)^k = (P_{M_r \cap M^\perp} \dots P_{M_1 \cap M^\perp})^k, \quad (32)$$

which leads to the following (not necessarily optimal) error bound:

$$\|(P_{M_r} \dots P_{M_1})^k(x) - P_M(x)\| \leq \|P_{M_r \cap M^\perp} \dots P_{M_1 \cap M^\perp}\|^k \|x\|. \quad (33)$$

If we assume that $\sum_{i=1}^r M_i^\perp$ is closed, then, by [2, Theorem 4.1], $\|P_{M_r \cap M^\perp} \dots P_{M_1 \cap M^\perp}\| < 1$ and the convergence is indeed linear. The above estimate can be found, for example, in [4, Remark 5.5.3], [11, Lemma 11.58] and [12, Lemma 9.2]. Note that for closed affine subspaces V_1, \dots, V_r with nonempty intersection $V := \bigcap_i V_i$ we can, similarly to (30), derive the equality

$$(P_{V_r} \dots P_{V_1})^k(x) - P_V(x) = (P_{V_r - V_r} \dots P_{V_1 - V_1})^k(x) - P_{V - V}(x - v) \quad (34)$$

for every $v \in V$. This leads to the following error bound:

$$\|(P_{V_r} \dots P_{V_1})^k(x) - P_V(x)\| \leq \left\| \prod_{i=1}^r P_{(V_i - V_i) \cap (V - V)^\perp} \right\|^k \|x - P_V(0)\|, \quad (35)$$

which yields linear convergence whenever $\sum_{i=1}^r (V_i - V_i)^\perp$ is closed. The above estimate can be found, for example, in [10, Theorem 6.6].

Funding. This research was supported in part by the Israel Science Foundation (Grant 389/12), the Fund for the Promotion of Research at the Technion and by the Technion General Research Fund.

References

- [1] N. ARONSZAJN, *Theory of reproducing kernels*, Trans. Amer. Math. Soc., 68 (1950), pp. 337–404.
- [2] C. BADEA, S. GRIVAUX, AND V. MÜLLER, *The rate of convergence in the method of alternating projections*, Algebra i Analiz, 23 (2011), pp. 1–30.
- [3] C. BADEA AND D. SEIFERT, *Ritt operators and convergence in the method of alternating projections*, J. Approx. Theory, 205 (2016), pp. 133–148.
- [4] H. H. BAUSCHKE, *Projection algorithms and monotone operators*, ProQuest LLC, Ann Arbor, MI, 1996. PhD thesis, Simon Fraser University (Canada).
- [5] H. H. BAUSCHKE, J. Y. BELLO CRUZ, T. T. A. NGHIA, H. M. PHAN, AND X. WANG, *Optimal rates of linear convergence of relaxed alternating projections and generalized Douglas-Rachford methods for two subspaces*, Numer. Algorithms, 73 (2016), pp. 33–76.
- [6] H. H. BAUSCHKE, F. DEUTSCH, AND H. HUNDAL, *Characterizing arbitrarily slow convergence in the method of alternating projections*, Int. Trans. Oper. Res., 16 (2009), pp. 413–425.
- [7] Y. CENSOR, W. CHEN, P. L. COMBETTES, R. DAVIDI, AND G. T. HERMAN, *On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints*, Comput. Optim. Appl., 51 (2012), pp. 1065–1088.
- [8] F. DEUTSCH, *Best approximation in inner product spaces*, vol. 7 of CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer-Verlag, New York, 2001.

- [9] F. DEUTSCH AND H. HUNDAL, *The rate of convergence for the cyclic projections algorithm. III. Regularity of convex sets*, J. Approx. Theory, 155 (2008), pp. 155–184.
- [10] ———, *Slow convergence of sequences of linear operators II: arbitrarily slow convergence*, J. Approx. Theory, 162 (2010), pp. 1717–1738.
- [11] ———, *Arbitrarily slow convergence of sequences of linear operators: a survey*, in Fixed-point algorithms for inverse problems in science and engineering, vol. 49 of Springer Optim. Appl., Springer, New York, 2011, pp. 213–242.
- [12] ———, *Arbitrarily slow convergence of sequences of linear operators*, in Infinite products of operators and their applications, vol. 636 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2015, pp. 93–120.
- [13] N. DUNFORD AND J. T. SCHWARTZ, *Linear operators. Part I*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.
- [14] I. HALPERIN, *The product of projection operators*, Acta Sci. Math. (Szeged), 23 (1962), pp. 96–99.
- [15] S. KAYALAR AND H. L. WEINERT, *Error bounds for the method of alternating projections*, Math. Control Signals Systems, 1 (1988), pp. 43–59.
- [16] M. L. LAPIDUS, *Generalization of the Trotter-Lie formula*, Integral Equations Operator Theory, 4 (1981), pp. 366–415.
- [17] G. PIERRA, *Decomposition through formalization in a product space*, Math. Programming, 28 (1984), pp. 96–115.
- [18] E. PUSTYLNİK, S. REICH, AND A. J. ZASLAVSKI, *Convergence of non-periodic infinite products of orthogonal projections and nonexpansive operators in Hilbert space*, J. Approx. Theory, 164 (2012), pp. 611–624.
- [19] ———, *Inner inclination of subspaces and infinite products of orthogonal projections*, J. Nonlinear Convex Anal., 14 (2013), pp. 423–436.
- [20] S. REICH, *A limit theorem for projections*, Linear and Multilinear Algebra, 13 (1983), pp. 281–290.
- [21] J. VON NEUMANN, *On rings of operators. Reduction theory*, Ann. of Math. (2), 50 (1949), pp. 401–485.